ST2334 Midterm Cheatsheet

github.com/reidenong/cheatsheets, AY23/24 S2

1. Probability and Counting

PIE:

For finite sets A, B and C,

$$|A \cup B| = |A| + |B| - |A \cap B|$$

$$|A \cup B \cup C| = |A| + |B| + |C| - |A \cap B| \\ -|A \cap C| - |B \cap C| + |A \cap B \cap C|$$

Conditional Probability:

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)}$$

$$P(A) = \frac{P(A \cap B)}{P(B \mid A)}$$

Inverse Probability Formula:

$$P(A \mid B) = \frac{P(B \mid A) \cdot P(A)}{P(B)}$$

Independent Events

Events A and B are independent ($\perp$) if and only if

- $P(A \mid B) = P(A)$ or $P(B \mid A) = P(B)$
- $\boldsymbol{P(A \cap B) = P(A) \cdot P(B)}$

Mutually Exclusive:

Events A and B are mutually exclusive if and only if $P(A \cap B) = \emptyset$

Law of Total Probability:

Suppose $A_1, A_2, ..., A_n$ is a partition of S. Then for any event B, we have

$$P(B) = \sum_{i=1}^{n} P(B \mid A_i) \cdot P(A_i)$$

Bayes' Theorem:

Suppose $A_1, A_2, ..., A_n$ is a partition of S. Then for any event B, we have

$$P(A_i \mid B) = \frac{P(B \mid A_i) \cdot P(A_i)}{P(B)}$$

2. Random Variables

Probability Mass Function:

For a *discrete* random variable X, the probability mass function (pmf) of X is

$$f(x) \begin{cases} P(X = x), & \text{for } x \in R_X \\ 0, & \text{for } x \notin R_X \end{cases}$$

(1) $f(x_i) \geq 0, \forall x_i \in R_X$
(2) $f(x) = 0, \forall x \notin R_X$
(3) $\sum_{x \in R_X} f(x) = 1$

Probability Density Function:

For a *continuous* random variable X:
(1) $f(x) \geq 0, \forall x \in R_X \wedge f(x) = 0 \forall x \notin R_X$
(2) $\int_{R_X} f(x)dx = 1$
(3) $f_X(x) \geq 0$, but not necessarily ≤ 1
(4) For some $a \leq b$,

$$P(a \leq X \leq b) = \int_a^b f(x)dx$$

Cumulative Distribution Function:

For *any* random variable X, the *cdf* of X is defined by

$$F(x) = P(X \leq x)$$

If X is a discrete random variable, then for any two numbers $a < b$, we have

$$P(a \leq X \leq b) = F(X \leq b) - F(X < a) \\ = F(b) - F(a-)$$

where $F(a-) = \lim_{x \uparrow a} F(x)$
= largest value in R_X that is less than a.

Further, $0 \leq f(x) \leq 1$

If X is a continuous random variable, then

$$F(x) = \int_{-\infty}^{x} f(t)dt,$$

$$f(x) = \frac{d}{dx}F(x)$$

$$P(a \leq X \leq b) = F(b) - F(a)$$

Further, $f(x) \geq 0$ but not necessarily ≤ 1.

CDFs are right continuous, have a maximum value of 1, and non decreasing.

Expectation and Variance

For a *discrete* random variable X, the expectation of X is

$$E(X) = \sum_{x \in R_X} x \cdot f(x).$$

For a *continuous* random variable X, the expectation of X is

$$E(X) = \int_{-\infty}^{\infty} x \cdot f(x)dx.$$

Properties:
(1) $E(aX + b) = aE(X) + b$
(2) $E(X + Y) = E(X) + E(Y)$
(3) Let g be an arbitrary function.
if X discrete,

$$E[g(X)] = \sum_{x \in R_X} g(x) \cdot f(x)$$

if X continuous,

$$E[g(X)] = \int_{R_X} g(x) \cdot f(x)dx$$

Variance:

$$\sigma_X^2 = V(X) = E\left[(X - \mu_X)^2\right]$$

If X is discrete, then

$$V(X) = \sum_{x \in R_X} (x - \mu_X)^2 \cdot f(x)$$

If X is continuous, then

$$V(X) = \int_{R_X} (x - \mu_X)^2 \cdot f(x)dx$$

Properties:
(1) $\forall X, V(X) \geq 0$. Equality holds when X is constant.
(2) $V(aX + b) = a^2 V(X)$
(3) $V(X) = E(X^2) - [E(X)]^2$
(4) The standard deviation of X is

$$\sigma_X = \sqrt{V(X)}$$

Discrete Joint Probability Function

Let (X, Y) be a 2-D discrete random variable. Its joint probability (mass) function is then given by

$$f(x, y) = P(X = x, Y = y), \forall (x, y) \in R_{X,Y}$$

Properties:
(1) $f(x, y) \geq 0, \forall (x, y) \in R_{X,Y}$
(2) $f(x, y) = 0$ if $(x, y) \notin R_{X,Y}$
(3) $\sum_{x \in R_X} \sum_{y \in R_Y} f(x, y) = 1$

Continuous Joint Probability Function

Let (X, Y) be a 2-D continuous random variable. Its joint probability (mass) function is then given by

$$f(x, y) = P(X \leq x, Y \leq y), \forall (x, y) \in R_{X,Y}$$

Properties:
(1) $f(x, y) \geq 0, \forall (x, y) \in R_{X,Y}$
(2) $f(x, y) = 0$ if $(x, y) \notin R_{X,Y}$
(3) $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y)dxdy = 1$

Marginal Probability Function

For a 2-D discrete random variable (X, Y) with joint probability function $f_{X,Y}$, the marginal probability function of X is as follows:

If Y is discrete, then for any x,

$$f_X(x) = \sum_{y \in R_Y} f_{X,Y}(x, y)$$

If Y is continuous, then for any x,

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x, y)dy$$

Conditional Distribution

For a 2-D discrete random variable (X, Y) with joint probability function $f_{X,Y}$, and marginal probability function for X as $f_{X(x)}$, the **conditional probability function of Y given $X = x$** is as follows:

$$f_{Y|X}(y \mid x) = P(Y = y \mid X = x) = \frac{f_{X,Y}(x, y)}{f_X(x)}$$

This can be interpreted as the distribution of Y given that $X = x$.

Independent Random Variables

Two random variables X and Y are independent if and only if for all x and y, we have

$$f_{X,Y}(x, y) = f_X(x) \cdot f_Y(y)$$

$R_{X,Y}$ needs to be a product space, ie. $R_{X,Y} = R_X \times R_Y$ for X and Y to be independent.

Properties:

Suppose X and Y are independent. Then

(1) For arbitrary subsets A and B,

$$P(X \in A; Y \in B) = P(X \in A) \cdot P(Y \in B)$$
$$P(X \leq x; Y \leq y) = P(X \leq x) \cdot P(Y \leq y)$$

(2) For arbitrary functions g and h, $g(X)$ and $h(Y)$ are independent.

(3) Independence is connected to conditional distributions.

$$\text{if } f_X(x) > 0, \text{ then } f_{Y|X}(y \mid x) = f_Y(y)$$
$$\text{if } f_Y(y) > 0, \text{ then } f_{X|Y}(x \mid y) = f_X(x)$$

Checking Independence

X and Y are independent if and only if for all x and y, we have

(a) $R_{X,Y}$, the range where the probability function is positive, is a product space. ie. the region $\{(x, y)\}$ is rectangular.

(b) $\forall (x, y) \in R_{X,Y}$,
$f_{X,Y}(x, y) = C \times g_1(x) \cdot g_2(y)$

Expectation and Variance of Random Variables

For any two variable function $g(x, y)$,

If (X, Y) is a 2-D discrete random variable, then

$$E[g(X, Y)] = \sum_x \sum_y g(x, y) \cdot f_{X,Y}(x, y)$$

If (X, Y) is a 2-D continuous random variable, then

$$E[g(X, Y)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x, y) \cdot f_{X,Y}(x, y) dx dy$$

Covariance

The covariance of two random variables X and Y is a measure of how the variables change together, defined as

$$\text{cov}(X, Y) = E[(X - \mu_X) \cdot (Y - \mu_Y)]$$

If (X, Y) is a 2-D discrete random variable, then

$$\text{cov}(X, Y) = \sum_x \sum_y (x - \mu_X)(y - \mu_Y) \cdot f_{X,Y}(x, y)$$

If (X, Y) is a 2-D continuous random variable, then

$$\text{cov}(X, Y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - \mu_X)(y - \mu_Y) \cdot f_{X,Y}(x, y) dx dy$$

Properties:

(1) $\text{cov}(X, Y) = E(XY) - \mu_X \mu_Y$

(2) if X, Y are independent, then $\text{cov}(X, Y) = 0$

(3) $\text{cov}(aX + b, cY + d) = ac \cdot \text{cov}(X, Y)$

(4) $V(aX + bY) = a^2 V(X) + b^2 V(Y) + 2ab \cdot \text{cov}(X, Y)$

Variance and Covariance

Using
$V(X + Y) = V(X) + V(Y) + 2\ \text{cov}(X, Y)$,
we can derive the following:

(1) For independent random variables X and Y,

$$V(X \pm Y) = V(X) + V(Y)$$

(2) For any random variables $X_1, X_2, ..., X_n$,

$$V\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} V(X_i) + 2 \sum_{i<j} \text{cov}(X_i, X_j)$$